ELSEVIER

Available online at www.sciencedirect.com

SCIENCE DIRECT

J. Differential Equations 196 (2004) 466–480

Journal of Differential Equations

http://www.elsevier.com/locate/jde

The approximate decomposition of exponential order of slow–fast motions in multifrequency systems

Fuzhong Cong[a,b]

[a] *School of Mathematics and Information Science, Shandong Institute of Business and Technology, Yantai 264005, People's Republic of China*

[b] *Office of Mathematics, Changchun Flight Academy of Air Force, Changchun 130022, People's Republic of China*

Received November 13, 2002; revised July 25, 2003

Abstract

This paper deals with multifrequency slow–fast systems. It is shown that, under a suitable change of coordinates, the system can be reduced to a simple form such that slow motions are described by autonomous equations except for exponential error of perturbations. Hence, the fast and slow motions are decoupled. The Newton rapid iteration is used. In addition, for a perturbation, only the smallness condition is needed.

MSC: 37J40

Keywords: Slow–fast system; Newton rapidly iteration; Approximate decomposition

1. Introduction and main results

Consider the slow–fast system

$$\dot{x} = \omega(y) + \varepsilon f(x, y, \varepsilon), \quad \dot{y} = \varepsilon g(x, y, \varepsilon), \tag{1.1}$$

where $x \in T^n = R^n/Z^n$ and $y \in D \subset R^m$, D is an open set, and $\varepsilon > 0$ is a small parameter.

E-mail address: congfz@263.net.

0022-0396/$ - see front matter
doi:10.1016/j.jde.2003.09.003

As well known, the above system (1.1) is important in the classical perturbation theory [2–4,6,8]. The variables x and y are said to be fast and slow, respectively. This paper discusses system (1.1) from the point of view of the decomposition of differential equations. To be more precise, we weaken the dependence of the right-hand sides of System (1.1) on fast variables. Under a series of changes of variables (1.1) is reduced to a simple form such that the slow motions can be described by an autonomous system except for an error of exponential order. Thus, the fast and slow variables are decoupled.

For some $\rho > 0$, we denote the ρ-neighborhood of $T^n \times D$ in $C^{n\times m}$ by $(T^n \times D) + \rho$. Define

$$\mathcal{O}(\tau,\gamma) = \{y : y \in D, |\langle \omega(y), k\rangle| \geqslant \gamma |k|^{-\tau}, k \in Z^n \backslash \{0\}\},$$

where (and below) for $k \in Z^n$, denote $|k| = |k_1| + |k_2| + \cdots + |k_n|$.

Theorem A. *For all $\varepsilon \in (0,1)$ assume that ω, f, and g are real analytic functions defined on $(T^n \times D) + \rho$. Let f and g be smooth in ε. For any positive constants τ and γ, if $y_0 \in \mathcal{O}(\tau,\gamma)$ and ε is small enough, then there exists a change of coordinates $\mathcal{S}_*$, defined on some neighborhood of y_0, such that $\mathcal{S}_*$ reduces* (1.1) *into*

$$\begin{aligned}\dot{X} &= \omega(y_0) + \Omega_*(Y,\sqrt{\varepsilon}) + \sqrt{\varepsilon} F_*^1(X,Y,\sqrt{\varepsilon}),\\ \dot{Y} &= \sqrt{\varepsilon}(G_*(Y,\sqrt{\varepsilon}) + G_*^1(X,Y,\sqrt{\varepsilon})),\end{aligned} \tag{1.2}$$

which satisfies the estimates

$$|\Omega_*(Y,\sqrt{\varepsilon})| < 2\alpha\sqrt{\varepsilon},$$

$$|G_*(Y,\sqrt{\varepsilon})| < 2\alpha,$$

$$|F_*^1(X,Y,\sqrt{\varepsilon})| + |G_*^1(X,Y,\sqrt{\varepsilon})| < \alpha \exp(-\beta \varepsilon^{-\frac{1}{2(n+\tau+2)}}),$$

where α and β are positive constants independent of ε.

Now we make some comments on Theorem A.

- If ω satisfies certain nondegeneracy condition on D, that is, the vector set consisting of ω and its derivatives suits some full rank condition, then, from the KAM theory [1,2,4,7,9,13,14], it follows that there exists a positive constant τ such that $\mathcal{O}(\tau,\gamma)$ is a Cantor set with positive measure, provided γ is not large. This shows that, in general, $\mathcal{O}(\tau,\gamma)$ is nonempty.
- We can approximately solve (1.1) by applying the equations

$$\dot{X} = \omega(y_0) + \Omega_*(Y,\sqrt{\varepsilon}), \quad \dot{Y} = \sqrt{\varepsilon} G_*(Y,\sqrt{\varepsilon}).$$

This accuracy achieves order $O(\exp(-\beta\varepsilon^{-\frac{1}{2(n+\tau+2)}}))$ of size ε of perturbation in a time $\exp(\frac{\beta}{2}\varepsilon^{-\frac{1}{2(n+\tau+2)}})$.

- As $n=1$, our situation is one of those in [4,10]. Of course, the results in [4,10] are global, but ours local.
- Treschev and other mathematicians researched the problem of separation of fast and slow motions by using averaging methods in a series of papers [10,12,15–18]. Recently, employing continuous averaging, Pronin and Treschev have accomplished the decomposition of fast and slow motions for system (1.1). In contrast to our result, they impose smallness condition upon the average values of f and g. In this sense, there is no restriction on the perturbation in Theorem A.
- Let $y_0 \in D$. If there exists $k_0 \in Z^n \backslash \{0\}$ that satisfies $\langle k_0, \omega(y_0)\rangle = 0$, then there is an integer matrix M with $\det M = 1$ such that $M\omega(y_0) = (\bar{\omega}(y_0), 0)^T$, where the vector $\bar{\omega} \in R^l, l < n$. Assume that $\bar{\omega}(y_0)$ satisfies the Diophantine condition

$$|\langle \bar{k}, \bar{\omega}(y_0)\rangle| \geqslant \gamma_0 |\bar{k}|^{-\tau_0}, \quad 0 \neq \bar{k} \in Z^l.$$

We introduce a transformation

$$y = Y', \quad x = M^{-1}\begin{pmatrix} X \\ Y'' \end{pmatrix}, \tag{1.3}$$

where $X \in T^l$, $Y = (Y', Y'')^T \in R^{m+n-l}$, which transforms (1.1) into the following:

$$\dot{X} = \bar{\omega}(y_0) + \varepsilon F(X, Y, \varepsilon), \quad \dot{Y} = \varepsilon G(X, Y, \varepsilon). \tag{1.4}$$

Here

$$\begin{pmatrix} F(X, Y, \varepsilon) \\ G(X, Y, \varepsilon) \end{pmatrix} = \begin{pmatrix} Mf\left(M^{-1}\begin{pmatrix} X \\ Y'' \end{pmatrix}, Y', \varepsilon\right) \\ g\left(M^{-1}\begin{pmatrix} X \\ Y'' \end{pmatrix}, Y', \varepsilon\right) \end{pmatrix}.$$

So we can apply Theorem A to (1.4). In [12], an analogous transformation has been used.

- Consider a nearly twisted mapping

$$\hat{x} = x + \omega(y) + \varepsilon f(x, y, \varepsilon), \quad \hat{y} = y + \varepsilon g(x, y, \varepsilon), \tag{1.5}$$

where $\omega(y)$ satisfies the usual Diophantine condition at y_0. By using the technique of iteration, similarly, we obtain the result.

Theorem B. *There exists a transformation of coordinates defined on some neighborhood of y_0 which changes* (1.5) *into*

$$\hat{X} = X + \omega(y_0) + \Omega_*(Y, \sqrt{\varepsilon}) + \sqrt{\varepsilon} f_*(X, Y, \sqrt{\varepsilon}),$$

$$\hat{Y} = Y + \sqrt{\varepsilon}(G_*(Y, \sqrt{\varepsilon}) + g_*(X, Y, \sqrt{\varepsilon})),$$

with

$$|\Omega_*| < c\sqrt{\varepsilon}, \quad |G_*| < c$$

and

$$|f_*| + |g_*| < c\exp(-c\varepsilon^{-\frac{1}{2(n+\tau+2)}})$$

for some positive constant c. Here τ *is the Diophantine exponent; that is,* τ *satisfies*

$$|\langle k, \omega(y_0)\rangle + k_0| \geqslant \gamma_0 |k|^{-\tau}, \quad \textit{for all } 0 \neq k \in Z^n \textit{ and } k_0 \in Z.$$

- For a nonautonomous differential equation

$$\dot{x} = \omega(y) + \varepsilon h(x, y, t, \varepsilon), \quad \dot{y} = \varepsilon l(x, y, t, \varepsilon), \tag{1.6}$$

where $(x, y) \in T^n \times D \subset R^n \times R^m$, and h and l are 1-periodic in t. Let $x_{n+1} = t$, and

$$X = \begin{pmatrix} x \\ x_{n+1} \end{pmatrix}, \quad H(X, y, \varepsilon) = \begin{pmatrix} h(x, y, t, \varepsilon) \\ 0 \end{pmatrix},$$

$$\tilde{\omega}(y) = \begin{pmatrix} \omega(y) \\ 1 \end{pmatrix}, \quad L(X, y, \varepsilon) = l(x, y, t, \varepsilon).$$

Then

$$\dot{X} = \tilde{\omega}(y) + \varepsilon H(X, y, \varepsilon), \quad \dot{y} = \varepsilon L(X, y, \varepsilon), \tag{1.7}$$

where $X \in T^{n+1}$. If $\omega(y)$ satisfies

$$\operatorname{rank}\left\{\frac{\partial^\alpha \omega}{\partial y^\alpha}:\ 1 \leqslant |\alpha| \leqslant \min\{n, m\}\right\} = n,$$

by [5,14], for $\tau > n^2 - 1$, the set

$$\tilde{\mathcal{O}}(\gamma, \tau) = \{y \in D : |\langle k, \tilde{\omega}(y)\rangle| \geqslant \gamma |k|^{-\tau}, 0 \neq k \in Z^{n+1}\}$$

is a Cantor set with positive measure. Hence, we can apply Theorem A to (1.7).
- Finally, consider a nonautonomous system

$$\dot{x} = \omega(y) + \varepsilon p(x, y, t, \varepsilon), \quad \dot{y} = \varepsilon q(x, y, t, \varepsilon), \tag{1.8}$$

where p and q are quasi-periodic functions with the same frequency ω^0, that is, $p(x, y, t, \varepsilon) = P(x, \omega^0 t, y, \varepsilon)$ and $q(x, y, t, \varepsilon) = Q(x, \omega^0 t, y, \varepsilon)$, and P and Q are defined on $T^{n+l} \times D \subset R^{n+l} \times R^m$, where l is the dimension of the vector ω^0. Let

$$X = \begin{pmatrix} x \\ \omega^0 t \end{pmatrix}, \quad \hat{\omega}(y) = \begin{pmatrix} \omega(y) \\ \omega^0 \end{pmatrix},$$

$$\tilde{P}(X, y, \varepsilon) = \begin{pmatrix} P(x, \omega^0 t, y, \varepsilon) \\ 0_l \end{pmatrix}, \quad \tilde{Q}(X, y, \varepsilon) = Q(x, \omega^0 t, y, \varepsilon),$$

where 0_l denotes the l-dimensional zero vector. Then (1.8) is reduced into

$$\dot{X} = \hat{\omega}(y) + \varepsilon \tilde{P}(X, y, \varepsilon), \quad \dot{y} = \varepsilon \tilde{Q}(X, y, \varepsilon). \tag{1.9}$$

Assume that ω^0 satisfies the Diophantine condition

$$|\langle k, \omega^0 \rangle| \geqslant \gamma |k|^{-\tau}, \quad 0 \neq k \in Z^l$$

for some positive γ and sufficiently large τ. Imposing some nondegeneracy upon $\omega(y)$, it can be shown that the set

$$\hat{\mathcal{O}}(\gamma, \tau) = \left\{ y \in D : |\langle k, \hat{\omega}(y) \rangle| \geqslant \gamma |k|^{-\tau}, 0 \neq k \in Z^{n+l} \right\}$$

is nonempty. Thus, we can use Theorem A to solve (1.8).

2. The slow–fast system with a fixed frequency

In this section we consider the slow–fast system

$$\dot{x} = \omega + \varepsilon f(x, y, \varepsilon), \quad \dot{y} = \varepsilon g(x, y, \varepsilon), \tag{2.1}$$

where f and g are real analytic on $(T^n \times D) + \rho$, and ω satisfies the Diophantine condition

$$|\langle k, \omega \rangle| \geqslant \gamma |k|^{-\tau}, \quad 0 \neq k \in Z^n,$$

for given positive constants γ and τ.

We denote the usual maximum norm on the determined domain by $|\cdot|$ under cases without any confusion. $[\cdot]$ denotes the integer part of a given real number. In the sequel, all constants $c_1, c_2, \cdots$ are positive and independent of ε. For any real analytic function $l(x, y, \varepsilon)$, we set

$$\bar{l}(y, \varepsilon) = \int_{T^n} l(x, y, \varepsilon) \, \mathrm{d}x.$$

Theorem C. *There exists a change of coordinates* $\mathcal{T}_* : (X, Y) \to (x, y)$ *defined on* $(T^n \times D) + \frac{1}{4}\rho$, *which is real analytic and reduces* (2.1) *to the form*

$$\dot{X} = \omega + \Omega_*(Y, \varepsilon) + \varepsilon f_*(X, Y, \varepsilon), \quad \dot{Y} = \varepsilon(G_*(Y, \varepsilon) + g_*(X, Y, \varepsilon))$$

with the estimate

$$|\Omega_*| < 2K_1\varepsilon, |G_*| < 2K_1, \quad |f_*| + |g_*| \leqslant K_1 \exp(-c_* \varepsilon^{-\frac{1}{n+[\tau]+2}})$$

for some positive constants K_1 *and* c_*.

In order to prove Theorem C by using Newton rapid iteration, we need the following lemma.

Lemma 1. *Assume that* $h(x)$ *is a real analytic function defined on* $T^n + (\rho + \delta)$ *with* $\rho > 0$ *and* $\delta > 0$. *Then the following equation*

$$\frac{\partial u}{\partial x}\omega = h(x) - \bar{h} \tag{2.2}$$

has a unique analytic solution with $\bar{u} = 0$. *Moreover, on* $T^n + \rho$, u *satisfies the estimate*

$$|u| \leqslant \alpha(\gamma, \tau)|h| \frac{1}{\delta^{n+[\tau]+1}},$$

where

$$\alpha(\gamma, \tau) = \frac{2^n}{\gamma}\left(\frac{n + [\tau] + 1}{e}\right)^{n+[\tau]+1} \sum_{j=1}^{\infty} \frac{1}{j^{2-\tau+[\tau]}}.$$

Proof. Expanding h as a Fourier series,

$$h(x) = \sum_{k \in Z^n} h_k e^{\sqrt{-1}\langle k, x \rangle}. \tag{2.3}$$

Obviously,

$$u = \sum_{0 \neq k \in Z^n} \frac{h_k}{\sqrt{-1}\langle k, \omega \rangle} e^{\sqrt{-1}\langle k, x \rangle}$$

is a solution of (2.2). Therefore, on $T^n + \rho$, by Cauchy's formula, we have

$$\begin{aligned}|u| &\leqslant \sum_{0\neq k\in Z^n} \frac{|h_k|}{\gamma|k|^{-\tau}} e^{\rho|k|} \\ &\leqslant \frac{|h|}{\gamma} \sum_{0\neq k\in Z^n} \frac{|k|^{\tau}}{e^{\delta|k|}} \\ &\leqslant \frac{2^n|h|}{\gamma} \sum_{j=1}^{\infty} \frac{j^{n+\tau-1}}{e^{j\delta}} \\ &\leqslant \frac{2^n}{\gamma}(n+[\tau]+1)^{n+[\tau]+1} e^{-n-[\tau]-1} \sum_{j=1}^{\infty} \frac{1}{j^{2-\tau+[\tau]}} |h| \frac{1}{\delta^{n+[\tau]+1}} \\ &\leqslant \alpha(\gamma,\tau)|h| \frac{1}{\delta^{n+[\tau]+1}}.\end{aligned}$$

Here the inequality

$$x^{n+[\tau]+1} e^{-x\delta} \leqslant \left(\frac{n+[\tau]+1}{\delta}\right)^{n+[\tau]+1} e^{-(n+[\tau]+1)}, \quad x\geqslant 0 \tag{2.4}$$

is used. The proof of (2.4) is obtained by finding the maximal value of the function $l(x) = x^{n+[\tau]+1} e^{-x\delta}$, $x\geqslant 0$. $\square$

Let

$$\delta = \varepsilon^{\frac{1}{n+[\tau]+2}},$$

$$N = \left[\frac{\rho}{8K\delta}\right] + 1,$$

$$D_i = (T^n \times D) + \left(\frac{3}{4}\rho - 2jK\delta\right), \quad j = 1, 2, \ldots, N,$$

$$\hat{D}_i = (T^n \times D) + \left(\frac{3}{4}\rho - (2j-1)K\delta\right), \quad j = 1, 2, \ldots, N,$$

where $K>0$ is a constant determined below. Theorem C will be proven by the inductive method. Assume that (2.1) is reduced to

$$\dot{x} = \omega + \Omega_i(y,\varepsilon) + \varepsilon f_i(x,y,\varepsilon), \quad \dot{y} = \varepsilon(G_i(y,\varepsilon) + g_i(x,y,\varepsilon)), \tag{2.5}$$

defined on D_i, where

$$|G_i| < 2K_1, \tag{2.6}$$

$$|\Omega_i| < 2K_1\varepsilon, \tag{2.7}$$

$$|f_i| + |g_i| \leqslant M_i = \frac{1}{2^i}K_1, \tag{2.8}$$

and

$$K_1 = \sup_{((T^n\times D)+\rho)\times(0,1)} |f(x,y,\varepsilon)| + \sup_{((T^n\times D)+\rho)\times(0,1)} |g(x,y,\varepsilon)| + 1.$$

Introducing a transformation $\mathscr{T}_i : (X, Y) \to (x, y)$,

$$x = X + \varepsilon u_i(X, Y, \varepsilon), \quad y = Y + \varepsilon v_i(X, Y, \varepsilon), \tag{2.9}$$

we require $\mathscr{T}_i$ to change (2.5) to the same form,

$$\dot{X} = \omega + \Omega_{i+1}(Y,\varepsilon) + \varepsilon f_{i+1}(Y, Y, \varepsilon), \quad \dot{y} = \varepsilon(G_{i+1}(Y,\varepsilon) + g_{i+1}(X, Y, \varepsilon)), \tag{2.10}$$

except the lower script. Take

$$\Omega_{i+1}(Y,\varepsilon) = \Omega_i(Y,\varepsilon) + \varepsilon\bar{f}_i(Y,\varepsilon), \tag{2.11}$$

$$G_{i+1}(Y,\varepsilon) = G_i(Y,\varepsilon) + \bar{g}_i(Y,\varepsilon). \tag{2.12}$$

Define $\mathscr{T}_i$ by the following equations:

$$\frac{\partial v_i}{\partial X}(X, Y, \varepsilon)\omega = g_i(X, Y, \varepsilon) - \bar{g}_i(Y,\varepsilon), \tag{2.13}$$

$$\frac{\partial u_i}{\partial X}(X, Y, \varepsilon)\omega = f_i(X, Y, \varepsilon) - \bar{f}_i(Y,\varepsilon) + \frac{\partial \Omega_i}{\partial Y}(Y,\varepsilon)v_i(X, Y, \varepsilon). \tag{2.14}$$

By Lemma 1, (2.13) and (2.14), on $\hat{D}_i$, we have

$$|v_i| \leqslant 2\alpha(\gamma,\tau)\frac{1}{\delta^{n+[\tau]+1}}|g_i|, \tag{2.15}$$

$$\begin{aligned} |u_i| &\leqslant \alpha(\gamma,\tau)\frac{1}{\delta^{n+[\tau]+1}}\left(2|f_i| + \left|\frac{\partial\Omega_i}{\partial Y}\right||v_i|\right) \\ &\leqslant c_0\alpha^2(\gamma,\tau)\frac{1}{\delta^{n+[\tau]+1}}(|f_i| + |g_i|), \end{aligned} \tag{2.16}$$

where the definition of δ and $K > 1$ are used.

From (2.5), (2.9), and (2.10), it follows that

$$\omega + \Omega_i(y,\varepsilon) + \varepsilon f_i(x, y, \varepsilon) = \dot{X} + \varepsilon\frac{\partial u_i}{\partial X}\dot{X} + \varepsilon\frac{\partial u_i}{\partial Y}\dot{Y},$$

$$\varepsilon(G_i(y,\varepsilon) + g_i(x, y, \varepsilon)) = \dot{Y} + \varepsilon\frac{\partial v_i}{\partial X}\dot{X} + \varepsilon\frac{\partial v_i}{\partial Y}\dot{Y},$$

from which (2.11)–(2.14) lead to the relations

$$
\begin{aligned}
\varepsilon\left(\mathrm{E}_n+\varepsilon\frac{\partial u_i}{\partial X}\right)f_{i+1}(X,Y,\varepsilon) &= \Omega_i(y,\varepsilon)-\Omega_i(Y,\varepsilon)-\varepsilon\frac{\partial \Omega_i}{\partial Y}v_i\\
&\quad +\ \varepsilon(f_i(x,y,\varepsilon)-f_i(X,Y,\varepsilon))\\
&\quad +\ \varepsilon\frac{\partial u_i}{\partial X}\Omega_{i+1}(Y,\varepsilon)\\
&\quad +\ \varepsilon^2\frac{\partial u_i}{\partial Y}G_{i+1}(Y,\varepsilon)\\
&\quad +\ \varepsilon^2\frac{\partial u_i}{\partial Y}g_{i+1}(X,Y,\varepsilon)\\
&\triangleq \varepsilon\left(\mathrm{E}_n+\varepsilon\frac{\partial u_i}{\partial X}\right)(F^1_{i+1}+F^2_{i+1}+F^3_{i+1}+F^4_{i+1})\\
&\quad +\ \varepsilon^2\frac{\partial u_i}{\partial Y}g_{i+1}(X,Y,\varepsilon)
\end{aligned}
\tag{2.17}
$$

$$
\begin{aligned}
\varepsilon\left(\mathrm{E}_m+\varepsilon\frac{\partial v_i}{\partial Y}\right)g_{i+1}(X,Y,\varepsilon) &= \varepsilon(G_i(y,\varepsilon)-G_i(Y,\varepsilon))\\
&\quad +\ \varepsilon(g_i(x,y,\varepsilon)-g_i(X,Y,\varepsilon))\\
&\quad +\ \varepsilon\frac{\partial v_i}{\partial X}\Omega_{i+1}(Y,\varepsilon)\\
&\quad +\ \varepsilon^2\frac{\partial v_i}{\partial Y}G_{i+1}(Y,\varepsilon)\\
&\quad +\ \varepsilon^2\frac{\partial v_i}{\partial X}f_{i+1}(X,Y,\varepsilon)\\
&\triangleq \varepsilon\left(\mathrm{E}_m+\varepsilon\frac{\partial v_i}{\partial Y}\right)(G^1_{i+1}+G^2_{i+1}+G^3_{i+1}+G^4_{i+1})\\
&\quad +\ \varepsilon^2\frac{\partial v_i}{\partial X}f_{i+1}(X,Y,\varepsilon).
\end{aligned}
\tag{2.18}
$$

Here u_i and v_i are functions in X and Y. In (2.17) and (2.18) we use the conclusion that matrices $\mathrm{E}_n+\varepsilon\frac{\partial u_i}{\partial X}$ and $\mathrm{E}_m+\varepsilon\frac{\partial v_i}{\partial Y}$ are invertible for sufficiently small ε, which should be proven below. On the basis of (2.15), (2.16), and the definition of δ, one has

$$|\varepsilon u_i| \leqslant c_0K_1\alpha^2(\gamma,\tau)\delta \leqslant K\delta, \tag{2.19}$$

$$|\varepsilon v_i| \leqslant 2K_1\alpha(\gamma,\tau)\delta \leqslant K\delta, \tag{2.20}$$

provided

$$K \geqslant \max\{c_0 K_1 \alpha^2(\gamma,\tau), 2K_1\alpha(\gamma,\tau)\}. \tag{2.21}$$

Eqs. (2.19) and (2.20) show that as $(X, Y) \in D_{i+1}$, $(x,y) \in \hat{D}_i$. In addition,

$$\max\left\{\left|\varepsilon\frac{\partial u_i}{\partial X}\right|, \left|\varepsilon\frac{\partial v_i}{\partial Y}\right|\right\} \leqslant c_1 K_1 \frac{1}{K} < \frac{1}{2}$$

if

$$K \geqslant 2c_1 K_1. \tag{2.22}$$

Here (2.15) and (2.16) are used. Therefore, $\mathrm{E}_n + \varepsilon\frac{\partial u_i}{\partial X}$ and $\mathrm{E}_m + \varepsilon\frac{\partial v_i}{\partial Y}$ are inverse, and

$$\max\left\{\left|\left(\mathrm{E}_n + \varepsilon\frac{\partial u_i}{\partial X}\right)^{-1}\right|, \left|\left(\mathrm{E}_m + \varepsilon\frac{\partial v_i}{\partial Y}\right)^{-1}\right|\right\} \leqslant c_2. \tag{2.23}$$

From the mean value theorem,

$$f_i(x,y,\varepsilon) - f_i(X,Y,\varepsilon) = \frac{\partial f_i}{\partial x}(\xi,\eta,\varepsilon)(x-X) + \frac{\partial f_i}{\partial y}(\xi,\eta,\varepsilon)(y-Y),$$

where $(\xi,\eta) = (tx,ty) + ((1-t)X,(1-t)Y)$, $0 \leqslant t \leqslant 1$. Thus, for $(X,Y) \in D_{i+1}$, $(\xi,\eta) \in \hat{D}_i$. By using Cauchy's formula, (2.9), (2.15), (2.16), and (2.23), we obtain that

$$\begin{aligned}|F^2_{i+1}| &\leqslant \varepsilon c_2\left(\left|\frac{\partial f_i}{\partial x}\right||u_i| + \left|\frac{\partial f_i}{\partial y}\right||v_i|\right)\\ &\leqslant c_3\frac{\varepsilon}{K\delta^{n+[\tau]+2}}(|f_i|+|g_i|) \leqslant \frac{c_3}{K}(|f_i|+|g_i|).\end{aligned} \tag{2.24}$$

By (2.9), (2.15), (2.16), (2.23), (2.11), (2.12), and Cauchy's formula, one has

$$\begin{aligned}|F^1_{i+1}| &\leqslant \varepsilon c_2\left|\frac{\partial^2\Omega_i}{\partial Y^2}\right||v_i|^2\\ &\leqslant \frac{c_4\varepsilon^2}{K^2\delta^{2(n+[\tau]+1)+2}}|g_i| \leqslant \frac{c_4}{K^2}(|f_i|+|g_i|),\end{aligned} \tag{2.25}$$

$$\begin{aligned}|F^3_{i+1}| &\leqslant \left|\frac{\partial u_i}{\partial X}\right|(|\Omega_i| + \varepsilon|f_i|)\\ &\leqslant 3K_1 c_5\frac{\varepsilon}{K\delta^{n+[\tau]+2}}(|f_i|+|g_i|) \leqslant \frac{3K_1 c_5}{K}(|f_i|+|g_i|),\end{aligned} \tag{2.26}$$

$$|F_{i+1}^4| \leqslant \varepsilon \left|\frac{\partial u_i}{\partial Y}\right| (|G_i| + |\bar{g}_i|)$$
$$\leqslant (2K_1 + K)\alpha(\gamma, \tau)\frac{\varepsilon}{K\delta^{n+[\tau]+2}}(|f_i| + |g_i|) \leqslant \frac{c_6}{K}(|f_i| + |g_i|). \tag{2.27}$$

Similarly to the estimates of f_{i+1}, we can derive

$$|G_{i+1}^1| \leqslant \varepsilon c_2 \left|\frac{\partial G_i}{\partial Y}\right| |v_i| \leqslant \frac{c_7 \varepsilon}{K\delta^{n+[\tau]+2}} |g_i| \leqslant \frac{c_7}{K}(|f_i| + |g_i|), \tag{2.28}$$

$$|G_{i+1}^2| \leqslant \varepsilon c_2 \left(\left|\frac{\partial g_i}{\partial x}\right| |u_i| + \left|\frac{\partial g_i}{\partial y}\right| |v_i| \right) \leqslant \frac{c_8}{K}(|f_i| + |g_i|), \tag{2.29}$$

$$|G_{i+1}^3| \leqslant \left|\frac{\partial v_i}{\partial X}\right| (|\Omega_i| + \varepsilon |f_i|) \leqslant \frac{3K_1 c_9}{K}(|f_i| + |g_i|), \tag{2.30}$$

$$|G_{i+1}^4| \leqslant \varepsilon \left|\frac{\partial v_i}{\partial Y}\right| (|G_i| + |\bar{g}_i|) \leqslant \frac{c_{10}}{K}(|f_i| + |g_i|). \tag{2.31}$$

Obviously,

$$\max\left\{ \varepsilon \left| \left(\mathrm{E}_n + \varepsilon \frac{\partial u_i}{\partial X} \right)^{-1} \frac{\partial u_i}{\partial Y} \right|, \varepsilon \left| \left(\mathrm{E}_m + \varepsilon \frac{\partial v}{\partial Y} \right)^{-1} \frac{\partial v_i}{\partial X} \right| \right\}$$
$$\leqslant c_{11} c_2 \varepsilon K_1 \frac{1}{K\delta^{n+[\tau]+2}} (|f_{i+1}| + |g_{i+1}|)$$
$$\leqslant \frac{c_{11} c_2 K_1}{K} (|f_{i+1}| + |g_{i+1}|). \tag{2.32}$$

Now we choose K such that (2.21), (2.22), and

$$K \geqslant \max\{16c_3, 16c_4, 48K_1c_5, 16c_6, 16c_7, 16c_8, 48K_1c_9, 16c_{10}, 4c_{11}c_2K_1\} \tag{2.33}$$

hold. Then, from (2.24)–(2.32), it follows that

$$|f_{i+1}| + |g_{i+1}| \leqslant \frac{1}{2}(|f_i| + |g_i|) \leqslant M_{i+1} = \frac{1}{2^{i+1}} K_1.$$

Inductively, on D_{i+1},

$$|G_{i+1}| \leqslant \sum_{j=0}^{i+1} |\bar{g}_j| < 2K_1, \quad |\Omega_{i+1}| \leqslant \varepsilon \sum_{j=0}^{i+1} |\bar{f}_j| < 2K_1\varepsilon.$$

We define a change of coordinates

$$\mathscr{T} = \mathscr{T}_0 \circ \mathscr{T}_1 \circ \cdots \mathscr{T}_N : D_N \to D_0$$

by the equations

$$x = X + \varepsilon u_*(X, Y, \varepsilon), \quad y = Y + \varepsilon v_*(X, Y, \varepsilon),$$

where $(T^n \times D) + \frac{1}{4}\rho \subset D_N = (T^n \times D) + (\frac{3}{4}\rho - 2NK\delta) \subset (T^n \times D) + \frac{1}{2}\rho$. By (2.19) and (2.20), we have

$$\max\{|\varepsilon u_*|, |\varepsilon v_*|\} \leqslant NK\delta \leqslant \frac{\rho}{8} + K\delta \leqslant \frac{1}{4}.$$

Here we need that $\rho < 1$ and δ is sufficiently small. Hence, $\mathscr{T}$ is invertible.

Let $f_* = f_N$, $g_* = g_N$, $G_* = G_N$ and $\Omega_* = \Omega_N$. Then, under the transformation $\mathscr{T}$, (2.1) is reduced to

$$\dot{x} = \omega + \Omega_*(y, \varepsilon) + \varepsilon f_*(x, y, \varepsilon), \quad \dot{y} = \varepsilon(G_*(y, \varepsilon) + g_*(x, y, \varepsilon))$$

on $(T^n \times D) + \frac{1}{4}\rho$, which satisfies

$$|f_*| + |g_*| \leqslant \frac{1}{2^N} K_1 \leqslant K_1 \exp(-c_{12}\varepsilon^{-\frac{1}{n+[\tau]+2}}), \quad |\Omega_*| < 2K_1\varepsilon, |G_*| < 2K_1. \tag{2.34}$$

The proof of Theorem C is finished. □

- If (2.1) is Hamiltonian, transformation (2.9) can be chosen as canonical. Hence, the final system in Theorem C is also Hamiltonian.
- The Diophantine condition can be weaken. In fact, by Arnold's technique of cutting functions [1], for sufficiently small and fixed $\varepsilon > 0$, we need to solve the equation

$$\frac{\partial u}{\partial x}\omega = \sum_{0 < |k| \leqslant L(\varepsilon)} f_k e^{\sqrt{-1}\langle k, x\rangle},$$

where $L(\varepsilon)$ is a given integer dependent on ε, which needs only ω satisfying the finite inequalities

$$|\langle k, \omega\rangle| \geqslant \gamma |k|^{-\tau}, \quad 0 < |k| \leqslant L(\varepsilon).$$

3. Proof of Theorem A

For any given $y_0 \in \mathcal{O}(\gamma, \tau)$ we introduce a transformation $\mathscr{S} : (X, Y) \to (x, y)$,

$$y = y_0 + \sqrt{\varepsilon} Y, \quad x = X.$$

Then

$$\dot{X} = \omega(y_0) + \sqrt{\varepsilon} F(X, Y, \sqrt{\varepsilon}), \quad \dot{Y} = \sqrt{\varepsilon} G(X, Y, \sqrt{\varepsilon}), \tag{3.1}$$

where

$$F(X, Y, \sqrt{\varepsilon}) = \frac{\omega(y_0 + \sqrt{\varepsilon} Y) - \omega(y_0)}{\sqrt{\varepsilon}} + f(X, y_0 + \sqrt{\varepsilon} Y, \varepsilon),$$
$$G(X, Y, \sqrt{\varepsilon}) = g(X, y_0 + \sqrt{\varepsilon} Y, \varepsilon).$$

Let

$$O = \{Y \in R^m \colon |Y| < 1\}.$$

For sufficiently small $\varepsilon > 0$, we discuss (3.1) on $(T^n \times O) + \rho_0$ with $\rho_0 > 0$. By Theorem C, there exists a real analytic transformation $\mathscr{T}_* : (X', Y') \to (X, Y)$ on $(T^n \times O) + \frac{1}{4}\rho_0$, which reduces (3.1) to the form

$$\dot{X}' = \omega(y_0) + \Omega_*(Y', \sqrt{\varepsilon}) + \sqrt{\varepsilon} F_*^1(X', Y', \sqrt{\varepsilon}),$$
$$\dot{Y}' = \sqrt{\varepsilon}(G_*(Y', \sqrt{\varepsilon}) + G_*^1(X', Y', \sqrt{\varepsilon})), \tag{3.2}$$

where Ω_*, G_*, F_*^1, and G_*^1 satisfy

$$|\Omega_*| < 2\alpha\sqrt{\varepsilon}, |G_*| < 2\alpha, |F_*'| + |G_*'| \leqslant \alpha \exp(-\beta \varepsilon^{-\frac{1}{2(n+[\tau]+2)}}),$$

for some positive constants α and β. So $\mathscr{S} \circ \mathscr{T}_*$ is just the change needed in Theorem A. $\square$

- We have to reduce (1.1) into a system with a fixed frequency before applying Theorem C. For this purpose, let

$$\omega(y) = \omega(y_0) + (\omega(y) - \omega(y_0)),$$

where $y_0 \in \mathcal{O}(\tau, \gamma)$ for some τ and γ. In order to ensure that $\omega(y) - \omega(y_0)$ is small enough, there is need to consider (1.1) on a small neighborhood of y_0. In addition, we also require that the system with the fixed frequency, reduced by us, is defined on a large domain of slow variables. So that neighborhood of y_0 is enlarged. The change of coordinates

$$y = y_0 + \varepsilon^\kappa Y, \quad x = X \tag{3.3}$$

suits the above conditions. Here $\kappa \in (0, 1)$. It is not difficult to find that $\kappa = \frac{1}{2}$ is optimal. In fact, inserting (3.3) into (3.1), we have

$$\dot{X} = \omega(y_0) + O(\varepsilon^\kappa) + \varepsilon f(X, y_0 + \varepsilon^\kappa Y, \varepsilon), \quad \dot{Y} = \varepsilon^{1-\kappa} g(X, y_0 + \varepsilon^\kappa Y, \varepsilon),$$

hence, $1 - \kappa = \kappa$, that is, $\kappa = \frac{1}{2}$. The new system has a minimal perturbation.
- From KAM theory we know that in formulae of the transformation of coordinates the denominators $\langle k, \omega(y) \rangle$ cannot vanish in the domain considered. This condition is fulfilled in single-frequency systems with nonvanishing

frequency and in systems with constant frequency that satisfy the Diophantine condition. Hence, in [10] and in Section 2 of this paper, the global results of the approximate decomposition of slow–fast motions are developed. However, for the general case, this condition fails. On the other hand, KAM theory also tell us that for $\tau > n-1$, $\bigcup_{\gamma>0} \mathcal{O}(\tau,\gamma)$ has full measure. Thus,we can require that $\langle k, \omega(y) \rangle$ does not vanish on some neighborhood of every $y_0 \in \mathcal{O}(\tau,\gamma)$, whose radius are order of $\sqrt{\varepsilon}$ (see [4, pp. 144, 162]). This is the geometric meaning of the square root of the original small perturbation parameter.

- To extend $\mathcal{S}_*$ in Theorem A in the entire Cantor set $\bigcup_{\gamma>0} \mathcal{O}(\tau,\gamma)$ is a complicate problem. This needs Whitney's analytic extensions theory. For details, see [11].

Acknowledgments

I am very grateful to the referees for helpful suggestions. I am thankful to Professor Jialin Hong for help. I thank the Postdoctor of Foundation of China and the Wang Kuancheng Foundation for partial support. This work was partially supported by Grant 10101030 from NSF–China and by a Grant from NSF–Shandong Province in China.

References

[1] V.I. Arnold, Proof of A. N. Kolmogorov's theorem on the preservation of quasi-periodic motions under small perturbations of the hamiltonian, Usp. Mat. USSR 18 (1963) 13–40.
[2] V.I. Arnold, Mathematical Methods of Classical Mechanics, Moskva, Nauka, 1974 (in Russian); English transl.: Springer, New York, Heidelberg, Berlin, 1978.
[3] V.I. Arnold, Geometrical Methods in the Theory of Ordinary Differential Equations, Moskva, Nauka, 1978 (in Russian); English transl.: Springer, New York, Heidelberg, Berlin, 1983.
[4] V.I. Arnold, V.V. Kozlov, A.I. Neishtadt, Mathematical Aspects of Classical and Celestial Mechanics, Springer, Beijing, 1993.
[5] F.Z. Cong, Y. Li, M.Y. Huang, Invariant tori for nearly twist mappings with intersection property, Chinese Northeast Math. J. 12 (3) (1996) 280–298.
[6] J.K. Hale, Ordinary Differential Equations, Wiley-Interscience, New York, 1969.
[7] A.N. Kolmogorov, On the conservation of conditionally periodic motions for a small change in Hamilton's function, Dokl. Akad. Nauk. SSSR 98 (1954) 525–530 (in Russian).
[8] P. Lochak, C. Meunier, Multiphase Averaging for Classical Systems, Springer, New York, Heidelberg, Berlin, 1988.
[9] J. Moser, On invariant curves of area preserving mappings of annulus, Nachr. Akad. Wiss. GöMath-Phys. Kl (1962) 1–20.
[10] A.I. Neishtadt, The separation of motions in systems with rapidly rotating phase, Prikl. Mat. Mekh. 48 (2) (1984) 197–204 (in Russian); English transl. J. Appl. Math. Mech. 48(2) (1984) 133–139.
[11] J. Pöschel, Integrability of Hamiltonian system on Cantor sets, Commun. Pure Appl. Math. 35 (1982) 653–695.
[12] A. Pronin, D. Treschev, Continuous averaging in multi-frequency slow–fast systems, Regular Chaotic Dyn. 5 (2000) 157–170.
[13] H. Rüssmann, On twist hamiltonians, talk on the Colloque, International: Mecanique Celeste et Systems Hamiltonians, Marseille, 1990.

[14] M.B. Sevryuk, Invariant tori of Rüssmann nondegenerate hamiltonian systems, Russian Math. Dokl. 346 (1996) 590–593.
[15] D. Treschev, An averaging method for Hamiltonian systems, exponentially close to integrable ones, Chaos 6 (1) (1996) 6–14.
[16] D. Treschev, The continuous averaging method in the problem of separation of fast and slow motions, Regular Chaotic Dyn. 2 (1997) 9–20 (in Russian).
[17] D. Treschev, Separatrix splitting for a pendulum with rapidly oscillating suspension point, Russian J. Math. Phys. 5 (1997) 63–98.
[18] D. Treschev, Continuous averaging in dynamical systems, Proceedings of the International Congress of Mathematicians, Beijing, Vol. III, 2002, pp. 383–392.